

WAIS Inc

Wide Area Information Servers

FOR IMMEDIATE RELEASE

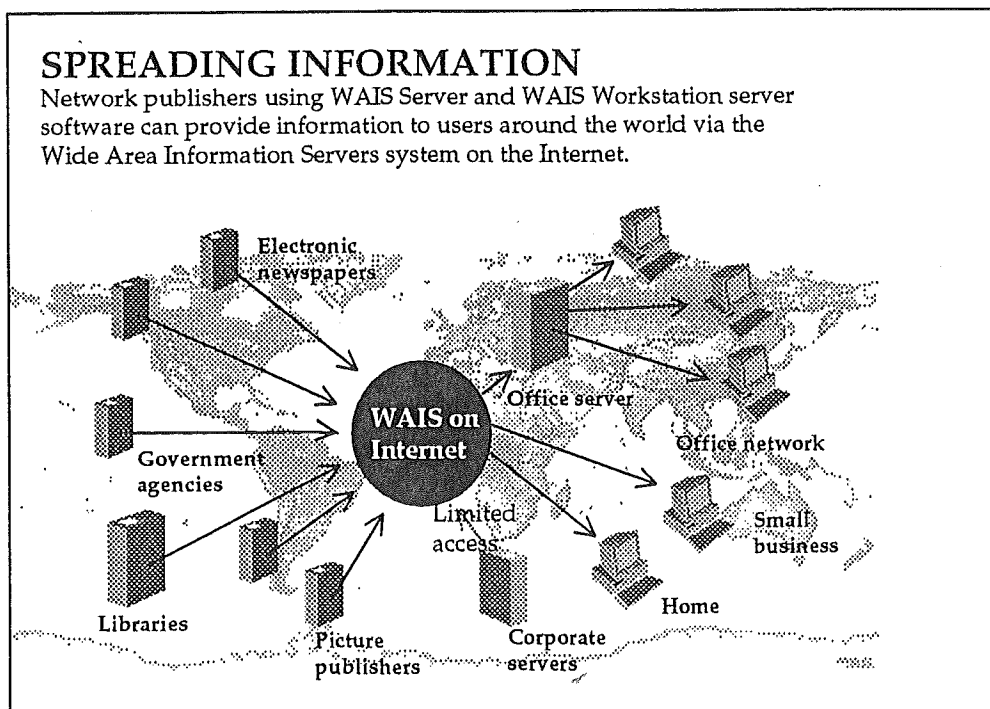
Contact: WAIS Inc.
info@wais.com
(415) 617-0444

WAIS INC. RELEASES NEW NETWORK PUBLISHING SOFTWARE

April 29, 1993, MENLO PARK, California – At a press briefing held today, Wide Area Information Servers Inc. (WAIS Inc.), a new company, announced its line of network publishing server software.

The products, WAIS Server™ for Unix and WAIS Workstation™ for Unix, are designed for commercial-quality network publishing applications. The software sells for \$10,000 to \$50,000.

"WAIS Inc. is helping publishers and companies reach the millions of people who are using today's wide-area networks," said Brewster Kahle, president of WAIS Inc. "We are combining a powerful search engine and a professional support organization with the WAIS system. The resulting system is creating a new breed of network publishers that can distribute information more broadly, quickly and inexpensively than previous technologies."



The WAIS™ system is an open network publishing system, based on international standards, that lets users easily find information anywhere on the network. The system, which has been tested globally on the Internet for two years, is used by people in more than 28 countries to search some 400 information sources around the world.

Mitchell Kapor of the Electronic Frontier Foundation and founder of Lotus Development Corp. said: "Millions of people already use the Internet. WAIS is an important tool helping people navigate through the vast oceans of information of the net, and WAIS Inc. is an important pioneer in building the tools which open new information markets."

WAIS Inc. customers include Perot Systems, the Library of Congress, Environment Protection Agency, Rice University and TRADE. WAIS Inc. is not at liberty to disclose the names of all customers.

WAIS Server is a powerful information server that is designed for client-server operation on wide-area networks and enterprise-wide LANs. Its features help users understand what a database contains and how it should be used. It also supports restricted access and logging for billing purposes. These logging features also provide maintainers with feedback on usage patterns.

The search engine at the core of the WAIS Server allows easy access to documents in their original formats – such as word processors, databases or newsfeeds. It answers natural language questions and uses relevance feedback (or "find me more like that one") features to help untrained users navigate through gigabytes. For more trained users, the WAIS Server answers Boolean queries and fielded searches, while still ranking the best answers highest.

WAIS Workstation offers all of the same features, but its use is limited to databases of 100 megabytes or smaller. This configuration is designed for enterprise information sharing and smaller network publishers.

For more information on WAIS Inc. products, call Nathaniel Lee at (415) 617-0444 or e-mail to info@wais.com.

#

WAIS™ is a trademark of Wide Area Information Servers Inc.

Internet and some innovative software are changing the nature of the library. Instead of being at the end of the block, a book collection will exist in cyberspace.

Good-bye, Dewey decimals

By David C. Churbuck

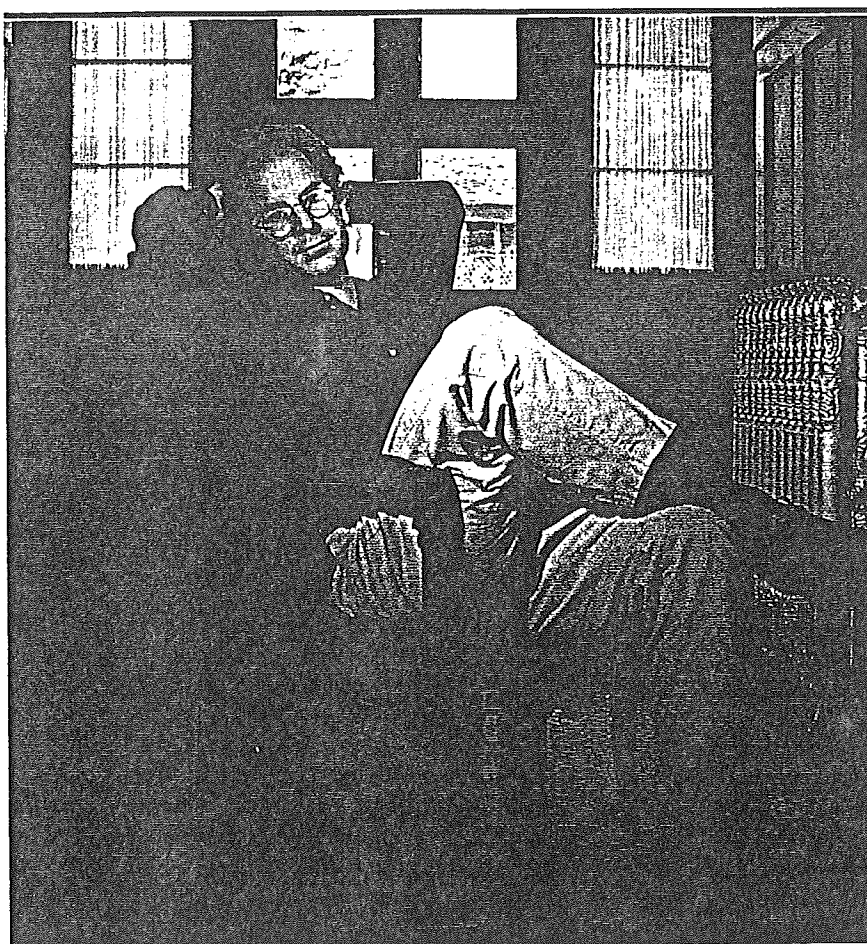
WAIS Inc. founder
Brewster Kahle
**Building the
electronic library
of the future.**

IT WOULD BE too much to argue that the jobs of the 152,000 librarians in the U.S. are in jeopardy. But it's fair to say that their jobs will change dramatically over the next two decades, courtesy of the Internet computer networking system and new software that controls Wide Area Information Servers, or WAIS.

Internet, a global network of computers formed by the government to connect universities, research labs and military complexes, was until recently largely limited to nonprofit users. But it is now taking on commercial customers (*FORBES*, July 8, 1991), who are finding the system an extremely economical way to gather and trade information. For as little as \$1 an hour, a subscriber to Internet can sit at a personal computer in his office, issue a request for information, and have the network route his request to libraries across the globe. The system will retrieve in a matter of seconds a collection of card-catalog citations that would have taken a lot of shoe leather to find in person. How else would a scholar in Chicago find out that "chaos" was touched on in a book written in 1741 in Latin and found in Harvard's library? Or what's in the Australian Defence Force Academy library. If he wants the volume, the researcher can follow up with a request for an interlibrary loan.

Full-text retrieval, still very limited at this point, is around the corner. When it comes, the local library as we know it all but disappears. In lieu of librarians we will have programmers and database experts.

One of the fathers of the Wide Area Information Servers concept is Brew-

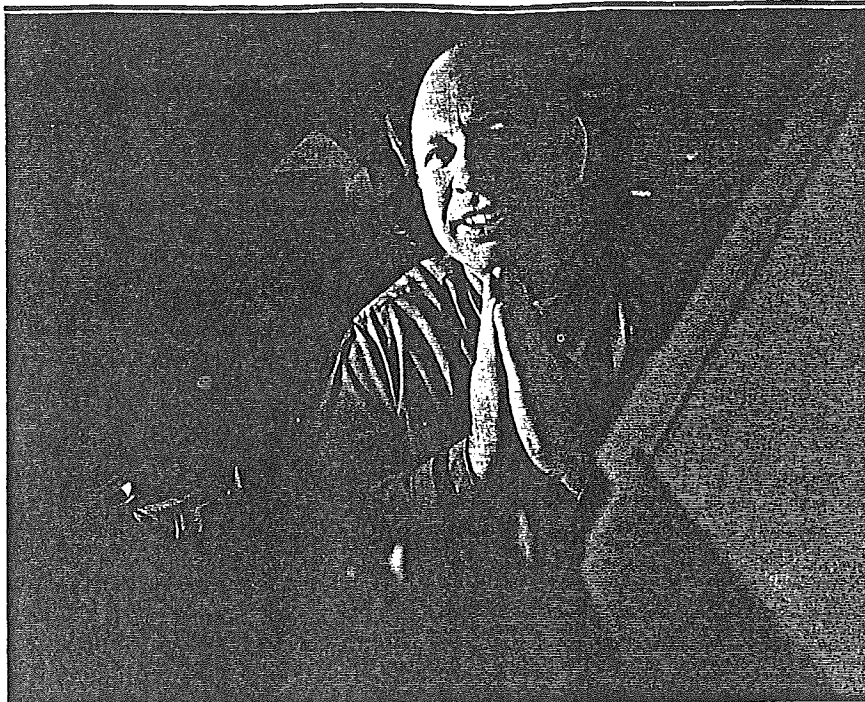


ster Kahle, who developed programs for searching databases when he worked at Thinking Machines Corp., the Cambridge, Mass. manufacturer of supercomputers. Thinking Machines is primarily interested in developing a market for its hardware, whose parallel processors are ideally suited to massive text searches. Kahle's interest is in the software.

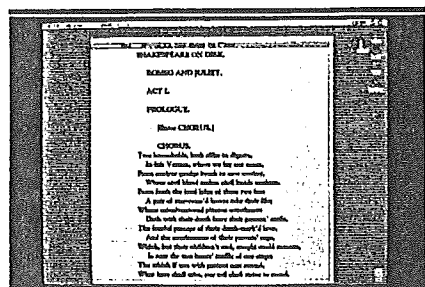
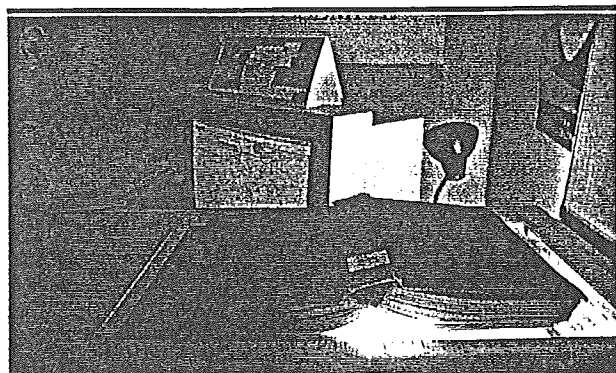
"People want to be able to pose a

question to the net[work] and not care where the answer comes from," says Kahle, 33, who recently left Thinking Machines to form WAIS Inc., in Menlo Park, Calif. Kahle's firm helps corporations and publishers set up electronic versions of copyrighted and public domain material and then make it accessible to users who are willing to pay for each download.

Kahle's software first paws through



Project Gutenberg founder Michael Hart; book scanning
Filling the shelves.



the raw text, indexing it to speed up later searches. Users can then submit plain-English queries. Ask for songs with the words "eyeball" and "toe jam," and it will come up with the Beatles' "Come Together." The system learns from trial and error, since the user can tell the computer which of the retrieved documents was closest to what he was looking for.

"WAIS is like a research librarian

who watches you read through a stack of information, taking notes on what you looked at first and set aside for future reference, and what information you threw away," explains Kahle. You can use WAIS to search a database on your own computer, but the software is optimized to handle the complexities of searches through physical distant databases.

Programmers at the University of Minnesota have taken WAIS a step further. To get information from a WAIS, you have to know exactly where that information is—which server has, say, a collection of song lyrics. The new software, called Gopher, makes it easier to navigate through the network. Richard Wiggins, Gopher coordinator at Michigan State University in East Lansing, Mich., explains the difference between WAIS and Gopher:

"WAIS is for direct searches. Gopher is a browsing tool." Enhancements to Gopher aim at Kahle's ideal—a query about the Beatles is answered with simultaneous searches through all the card catalogs on the network.

More libraries are joining the two-year-old Gopher system all the time, but the best is yet to come, including the 22-million-volume Library of Congress.

As with any new technology, some librarians are finding it hard to adapt. Says Wiggins: "In general terms, this technology is of interest to libraries, but it is frightening at the same time, setting off a struggle between academics who want knowledge spread around and librarians who want to control it."

Wide area information fanatics dream of an interlibrary loan system on a grand scale, in which only one copy of each book or magazine is in the system at all. That one copy could be used by many people at the same time and it would never be lost or overdue. Also, to the consternation of publishers thinking about royalties, it could be easily duplicated.

For this system of the future to become real, existing libraries of printed books will have to be digitized. That's the objective of Project Gutenberg, the creation of Michael Hart, professor of electronic text at Illinois Benedictine College in Lisle, Ill. If Kahle is building the shelves for the electronic library of the future, Hart is filling them. Gutenberg transcribes books, mostly in the public domain, for free digital distribution, whether via floppies or Internet.

For most of the 22 years that Hart has been at work on Gutenberg, digitizing has meant the slow and arduous retyping of the classics. Now, however, all Hart has to do is shred the binding off a book and feed the sheets into a scanner. Gutenberg has 50 titles on-line so far and expects to have 10,000 ready by the year 2001. That's nothing next to what the Library of Congress could crank out if it got serious about digitizing its collection, a project just getting under way. And if the programmers can solve some knotty problems involving pay-per-view and copying, there's no reason that newly published books cannot go on the system, too.

Gold Rush

SUDDENLY, A DIGITAL GOLD RUSH IS SWEEPING AMERICA.

The sprawling Internet, growing by one million users per month, remains raw and disorganized. And while nobody could argue that it's ready for prime time, the network's hour has surely arrived for corporations seeking low-cost connections to customers, prospects, and far-flung employees. What sprang to life nearly two decades ago as a research-only network has been freed from the red tape that hindered its commercial use. Technology companies have been the first to master its quirks in the hopes of generating profits and finding a competitive edge. Now, it's hard to find a company that isn't willing to give the Internet a chance.

What are the new uses for the Internet? Some companies are relying on it to connect remote offices as a kind of virtual local-area network (LAN). Others are creating their own Internets, inexpensive wide-area networks (WANs) that increasingly have all the security and tools of proprietary networks at a fraction of the cost. Others are discovering what has become, in effect, a new market: They're outsourcing revenue-creating products such as databases or catalogs on the Internet. So a publisher, for example, may hire a service provider to post its catalog of books online, thus making it available to 15 million users.

Then there are those who use the Internet as a step beyond the telephone or fax machine for customer support; for example, to transmit code fixes to customers who need immediate help. There are advertising agencies that send not only ad page mock-ups over the Internet but also compressed video images and sound bites to clients across town, speeding up the collaborative process (not to mention putting armies of bike messengers out of work).

But even Internet evangelists admit that everyone doesn't yet realize how to tap the Internet's potential. Part of the problem is that corporate decision makers often have never actually seen what's available, much less used file transfer protocol (FTP) to log onto a remote computer. Complains Eric Dentler, manager of sales systems at Nihon Sun, the Japanese sub-

siary of Sun Microsystems Inc.: "Until you can tell the CFO [chief financial officer] how they can make money [on the Internet], they don't move forward." And for some companies thinking of selling goods or services online, there are such tricky matters as figuring out how much to charge and how to collect from people who buy over the Internet.

Not Something You Just Tap Into

The Internet is the ultimate laissez-faire marketplace. Think of it as a town square that is one part farmers market, one part town hall, and one part post office. But unlike other networks that support commerce, it lacks a police department. Traffic, for the most part, flows chaotically. This free-flowing anarchy is what has led to the network's phenomenal growth, but it is also what is so confusing to novice users.

And unlike online services that provide corporations with crucial competitive research, such as Dow Jones & Co. Inc. and Mead Data Central Co.'s Dialog, the Internet's value isn't stored in databases. "What is there out there on the Internet for businesses in terms of real content?" asks Bill Washburn, executive director of the Commercial Internet Exchange (CIX), a trade group that has made it easier for companies to conduct business on the Internet. "I think that's a question a lot of people aren't owning up to."

Gigabytes have been spewed pontificating about the Internet's possibilities. But how many businesses really need National Weather Service satellite photos or access to the KGB's once-secret files? The content, or software of the Internet, still lacks a business focus.

"Netnews (the Internet's homegrown bulletin boards on a plethora of topics) is great for some people," says Washburn. "But for others it's a bunch of flame wars. Who are we kidding here?"

Nor are the Internet's current attractions easily accessible. "The Internet is not something you just tap into," explains Tod Dagres, a vice president of the Yankee Group, adding that companies hoping to use it commercially should expect to provide

**The possibilities are
virtually endless,
the price—at least for
the moment—is right,
so companies are
discovering the
rewards to be reaped
by hooking up to the
Internet. Of course,
not everybody is
ready to cash in**

By Jonathan Littman

ome training and education. Even Internet applications that should be relatively intuitive—such as sending electronic mail and files—require training and, yes, management.

The current commercial free-for-all began in 1991 when the federal government recognized the inevitable and made it known that it no longer intended to limit the network's backbone for use in research. That policy change created the incentive for

telephone charges, Alain Pinel provided its agents with Internet mail. Agents accustomed to sending mail internally over the Los Gatos office's Next local-area TCP/IP network could now send e-mail to customers and partners.

"It provides the ability to access our clients 24 hours a day," says Helen Pastorino, president of the real estate firm. Agents send listings, market analysis, and contingency documents

over the net to clients in Silicon Valley, the East Coast, Australia, and England.

Not satisfied with merely communicating with its customers, Alain Pinel created a cheap but effective wide-area Internet

with its partners. Agents send escrow information back and forth to local title companies over the Internet. Eager for Alain Pinel's business, seven title companies have made their own inexpensive dial-up arrangements to be on the Internet at a fraction of what electronic data interchange (EDI) costs. The title companies pay \$50 a month for Internet dial-up access (messages are sent each hour), plus local phone charges.

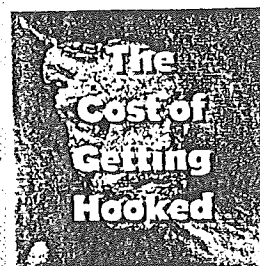
"We're just in our infancy, we're just starting to get intra-company usage," says Mark Richards, a partner in Alain Pinel, who has watched the start-up blossom in three years to four offices with 200 agents, 30 support-staff personnel, 800 listings, and annual home sales of \$600 million. "I can see our Internet use growing. We haven't really connected the mortgage companies, all the inspectors, the termite companies, the geological people, the roof report."

Many companies are finding that the Internet provides medium- to high-tech services they increasingly depend on. When Mary Cronin, author of "Doing Business on the Internet," interviewed a hundred Internet-connected corporations, she was surprised by how many manufacturing companies are heavy Internet users. "These small companies doing manufacturing are really finding it to be a resource," says Cronin, who found companies are using the net to receive bug fixes from vendors and helpful users. "They're not isolated anymore."

Engineers, scientists, and researchers in pharmaceutical, biotech, chemical, and oil companies often share tips and information on the net. Sometimes connections are made that lead to partnerships or joint ventures between professionals or companies.

Qualcomm Inc., a San Diego telecommunications company, receives résumés over the Internet to meet its expanding hiring needs. The résumés flow into a Resumex application on a Sun SPARCstation IPX, automatically matching the inquiries to the appropriate job openings and hiring managers. Last summer the firm began running ads in trade magazines in which the headline was its Internet address: Job@qualcomm.com.

For some novice users, finding information on the Internet is the equivalent of



Dial-up service for a few dozen employees can cost as little as \$100 a month. Sprintlink charges \$16,750 a year for a 56-kilobyte connection and the lease on a router (plus minimal local phone costs), often ample for modest use by several hundred employees. T1 connections (1.544 megabytes) offer nearly 30 times the volume of traffic—generally enough for a few thousand users—for about double the annual cost: \$30,000 to \$35,000.

the major Internet providers—Performance Systems International Inc. (PSI), Uunet Technologies Inc., and General Atomics' Cerfnet—to create their own commercial backbones that skirt the National Science Foundation's NSFNet. These providers, along with nine other providers such as Sprintlink and Narnet, formed CIX. Today, virtually every form of commercial traffic is acceptable through CIX providers except, of course, universally frowned-on unsolicited junk mailings.

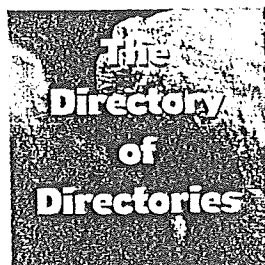
The potential for commercial exploitation was always there. First, the Internet's a TCP/IP protocol network, making it no stranger to most large corporations. Because many already have TCP/IP networks in place, it's relatively simple to connect offices in distant locations over the Internet. Second, low connection costs provide an alternative to building and maintaining expensive proprietary WANs (which also require leasing expensive long-distance lines) or joining limited commercial networks like Compuserve and America Online. Because companies only pay for access to the Internet and local phone charges, costs per user are often extremely low.

Third, the Internet is not only far larger than online commercial services, such as Compuserve, it also provides a wealth of other resources while still enabling users to send and receive e-mail to those smaller networks. By the time you read this article, roughly one million "host" machines in more than 137 countries will be connected to the Internet. Assuming about 10 users per host and a *monthly* growth rate of seven percent, the total number of users should approach 20 million by the end of 1994. According to the Internet Society, 30 percent of the Fortune 500 is on the Internet, and two-thirds of Internet users work for major corporations.

Business Digs It

Plain old business communication with customers remains one of the most basic uses of the Internet. Three years ago Alain Pinel Co., then a fledgling real estate firm in Los Gatos, Calif., tried to get an edge on the depressed Silicon Valley real estate market. But the company found it difficult to stay in touch with the overworked engineers, marketing types, and sales people who make up their clientele.

Enter the Internet. For a \$180 dedicated (continuous access) monthly dial-up fee to Portal Communications Co. of Cupertino, Calif., plus local



Since April, AT&T has been collecting "white pages" references for information servers, data and software archives, computing centers, and network providers. Access to the white and yellow Directory of Directories is free to Internet users. Only those listing their databases or services in the yellow pages must pay a fee.

Commerce on the Internet

searching for a needle in a haystack. New navigation and resource tools such asarchie, gopher, WAIS, and Worldwide Web are making it easier for companies to publish, index, search, and retrieve business resources on the net. The National Science Foundation recently hired AT&T Business Communications to create a "white" and commercial "yellow" pages for the Internet, as part of its InterNIC program, a five-year project designed to provide public information about the Internet. Needless to say, this type of service is fundamental for business networking.

Global Opportunities

Consultants and start-ups are trying to peddle the Internet, with its generally upwardly mobile user base, as an inexpensive yet effective marketing tool. A single example: some are partnering with companies to offer such new services as the ordering of magazines, books, or technical articles.

Last summer the publishers of the *New Republic* and a new marketing firm called the Internet Co. joined forces to create The Electronic Newsstand. This "electronic" magazine rack enables Internet users to browse the table of contents or search selected articles from magazines such as the *New Yorker*, the *Economist*, and the *New Republic*.

At the August 1993 Interop conference, O'Reilly & Associates, a Sebastopol, Calif. publishing company, introduced the Global Network Navigator (GNN), an Internet resource center in which users can "browse product brochures, press releases, and white papers; download demonstration software; or interact with a company representative," according to officials. Companies can submit hypertext advertisements that run in this GNN marketplace or in a general GNN guide to the Internet.

Less publicized are the scores of technical publishers that are making everything from technical articles to manuals and documentation available on the net—for a price. Those still squeamish about accepting credit card numbers over the net simply list toll-free 800 telephone numbers.

Others, for a surprisingly low fee, will place your company's inventory or catalog of merchandise, whatever it might be, on the Internet. "In some sense our service is like outsourcing your database," says Subu Subramanian, district manager of AT&T's business-communications services. "The [client] company doesn't have the ongoing responsibility. It doesn't have to incur the overhead of setting up a network connection." AT&T charges \$160 per month for posting databases in the 25- to 50-megabyte range, not counting security and access controls.

One competitor, Enterprise Integration Technology (EIT) of Palo Alto, Calif., will put a company's 30-page quarterly catalog on the Internet—potentially reaching millions of customers—for about \$15,000 a year. EIT customizes these catalogs to maximize



Search
and
you will
find

Search and menu programs likearchie and gopher help users find, browse, and transfer data widely dispersed on the net. Popular for "publishing" databases on the net, WAIS, the wide-area information server, enables users to build scripts of interest that can be run daily to create a sort of individual electronic newspaper. Worldwide Web provides the ability to link related words, pictures, and ideas around the Internet in a hypertext environment. Finally, recently introduced commercial versions of these public-domain tools answer corporate concerns of consistency, training, documentation, and support.

the multimedia possibilities of the Internet, using techniques such as hypertext, which is not currently available on other commercial networks. EIT, like AT&T and other information providers, offers a complete off-site, turnkey service, including its own servers, software, and T1 Internet connection.

It doesn't take a rocket scientist to see where this trend is headed: commercial overload. The multimedia capabilities, attractive demographics, and cheap access of the Internet will inevitably turn it into an advertising vehicle. "If Lexus, the car company, can represent their products [on the Internet], and create a place where customers can read their technical notices," predicts Rob Raisch, president of the Internet Co., "we suddenly provide an environment where consumers can make appropriate decisions." Of course, the notion makes old-time net hackers a trifle nauseous.

The Price Is Right

But even the Internet's heaviest users have yet to tap its full potential. At Unocal in Los Angeles, Internet usage is flourishing despite the fact that the company isn't particularly pushing it. "We haven't done anything," says Peter Ho, network systems engineer. That hasn't stopped intrepid Unocal engineers from traversing the Internet to access seismic data, satellite maps, and land surveys to search for oil and geothermal energy prospects. They also read select newsgroups on the Internet, and send and receive mail and files with other researchers and universities. Ho says one of the most practical uses is to receive new releases, bug fixes, patches, and online technical support from software and hardware vendors. Ho adds two Unocal users a week to the Internet, and the company spends about \$12,500 a year for a 56-kilobyte connection to the Cerfnet regional network and a leased line.

While Unocal has eschewed formal training or the development of Internet-specific tools, giant Lockheed Corp. has created a company-wide set of resources for its Internet travelers. Survival is its motive. Faced with cuts in defense spending, the defense contractor has had to watch costs and find new opportunities. Two years ago, Lockheed, already a veteran Internet e-mail user, decided to "create an electronic infrastructure so that researchers could share information, synergize, cooperate, [and] collaborate on research and projects for customers," says Michael Carroll, manager of advanced computer and software applications at corporate headquarters in Calabasas, Calif.

Several architectures were evaluated. But Carroll says the defense contractor "honed in on Internet-like communication" because they "already had a TCP/IP backbone in place. It was fairly straightforward to expand." Public-domain versions of client-server products such as WAIS, an Internet full-text information-indexing and retrieval program, made it inexpensive and easy to build Lockheed's internal Internet databases. "The price was right," says Carroll. "We had the ability to try [the Internet] without significant investment."

Carroll pooled together a few thousand hours of volunteer work and created the Technology Broker System (TBS), a hypermedia search and retrieval environment that bridges the Internet. Employees at both Lockheed Missiles in Space in